This is a repository copy of *Pareto-improving policies for an idealized two-zone city served by two congestible modes*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/120550/

Version: Accepted Version

INTERNATIONAL SYMPOSIUM ON TRAFFIC AND TRANSPORTATION THEORY

ISTTT22

Pareto-improving policies for an idealized two-zone city served by two congestible modes

Shu-Xian Xu [a,b], Ronghui Liu [b] *, Tian-Liang Liu [a] and Hai-Jun Huang [a]

[a] School of Economics and Management, Beihang University, Beijing 100191, China

[b] Institute for Transport Studies, University of Leeds, Leeds LS2 9JT, UK

Abstract

We study urban structure and traffic congestion of a monocentric city by idealizing its suburb and its core as two zones and then exploring what would happen when they are connected by a congestible highway and a crowded railway system. We introduce dynamic congestion effect into commuters' departure-time and mode choice behaviours, and analyse the endogenous interactions between their travel and residential relocation choices. Studies ignoring dynamic departure-time behaviour show an ambiguous effect of transit improvements to the city. However, we find that transit improvement has a definitive impact on city structure: it increases the residents' equilibrium utility, at a cost of increased suburb land use. We show that it is possible to design Pareto-improving land-use and transit policies which benefit the residents without causing urban sprawl. We provide analytically the existence conditions of such policies and suggest that a high return of land use tax to subsidize transit improvement is required.
Peer-review under responsibility of the scientific committee of the 22nd International Symposium on Transportation and Traffic Theory.

Keywords: Monocentric city model; Dynamic congestion; Land use tax; Transit subsidy; Urban sprawl; Pareto improving.

* Corresponding author. Tel: +44 113 343 5338; fax: +44 113 3435334.
  E-mail address: R.Liu@its.leeds.ac.uk

1. Introduction

A city's transportation system is intrinsically linked to its geographic structure and its economic activities. Rapid urbanization seen in cities around the world has led to investments in rail transit systems as ways to alleviate traffic congestion. On the one hand, the developments of such systems and the accompanying traffic management policies (e.g. pricing, transit subsidy) have a direct impact on where people choose to live and how they travel to work (McDonald and Osuji, 1995; Li et al., 2012; Ma and Lo, 2012). On the other hand, residential locations influence and determine the commuters' travel decisions (e.g. on travel mode, departure time, etc) (Boyce and Southworth, 1979). Thus, understanding the interplays between urban economic activities and transportation developments, and the role rail transit plays on travel behaviour and on urban land use, is important in shaping the city's development.

The standard model of urban structure and urban traffic congestion provide a powerful framework to explore these complex interrelationships, and the most basic form is the monocentric city model where all jobs are located in the Central Business District (CBD). The model describes the spatial structure of the city as the result of a trade-off between land rents and commuting costs (Alonso, 1964; Mills, 1967; Muth, 1969). The differences in commuting cost across the city are compensated by the differences in land price or housing price. This compensatory price variation, which reconciles suburban residents to long and costly commuting trips, is shown to have a profound effect on the spatial structure of the city.

When a public transit system is introduced to a city, it naturally alters the transportation landscape of the city and the cost of travel, which in turn affects the spatial structure of the city. Capozza (1973) was the first to integrate public transit into urban economic analysis and established a general spatial equilibrium model in a monocentric city with two transportation modes. He found that the addition of a subway system to a city reduces the overall transportation costs and the city size. Building on Capozza's work, Sasaki (1989) showed that reducing the public transport cost also leads to a contraction in city size. Su and DeSalvo (2008) further identified an inverse relation between transit subsidy and urban sprawl, and a direct relation between auto subsidy and urban sprawl. A common limitation however in the above studies of the two-mode city is that they ignored the transport congestion effects.

A feature in modern cities is the peak-hour traffic congestion. Modelling congestion in an urban economics setting, where both commuters' spatial distribution and travel behaviour (in their mode and departure time choices) are endogenous, raises challenges to urban economic analysis. Most existing cities models tend to assume a static travel cost function which varies only with distance. Strotz (1965) was the first to raise the congestion effect on the economy of urban communities. This is followed by studies of city models with travel costs which vary not only with distance but also with traffic density (e.g. Solow, 1972; Anas and Xu, 1999; Li et al. 2012). Congestion tolls are usually proposed as an effective way to internalize congestion externality, and their effect on a monocentric city is found to centralize the population towards the city centre (e.g. Wheaton, 1998; Verhoef, 2005). These studies, however, are concerned with monocentric cities served by only a single travel mode.

Few by far have examined congestion externality on urban spatial equilibrium with alternative transport modes, with the exceptions of Haring et al. (1976) and Buyukeren and Hiramatsu (2016). Haring et al. (1976) showed that increasing public transport capacity reduces the land rent differentials between the CBD and the city fringe, and lowers the equilibrium commuting costs. Using a discrete core-suburb monocentric model, Buyukeren and Hiramatsu (2016) found that modal substitution effect can limit the centralizing force of anti-congestion policies (such as road pricing) and lead to the acceleration of urban sprawl. In the above studies, the dynamic congestion effect on commuters' departure time choice is not considered.

In reality, congestion phenomena is highly dynamic and sensitive to traffic flow levels. Likewise, commuters are sensitive to congestion levels and they may choose to use the less congested mode to travel and/or to choose to depart early or late in order to avoid congestion. Thus, their generalized travel cost is determined not only by which mode they choose, but also by their trip timing and schedule delay costs. The classic bottleneck model first proposed by Vickrey (1969) provides a framework for analysing commuters' departure time choices as they vary endogenously with the dynamic nature of congestion. The bottleneck model has also been employed to examine the dynamic congestion effect on a combined residential location and departure time choice. Arnott (1998) incorporated

the Vickrey's bottleneck model into a model of a discrete core-suburb monocentric city connected by a road with a bottleneck. Contrary to the standard urban economics model which suggests that congestion tolling results in a more concentrated city, Arnott showed that when departure time decisions are considered, congestion tolling may have less pronounced effects on urban structure. In fact, imposing an optimal congestion toll without redistributing its revenues will have no effect on commuting costs and residential locations of the commuters. Using a spatially continuous monocentric city with a bottleneck at the entrance to the CBD, Gubins and Verhoef (2014) found that congestion tolling causes residents to spend more time at home and to have larger houses, thereby leading to urban sprawl. Taking commuters' heterogeneity into account in a continuous monocentric city model, Takayama and Kuwahara (2016) examined the influence of time-varying congestion tolling and showed that congestion tolling leads to increased population in the suburb and generates urban sprawl. The above studies of dynamic congestion effect on cities, however, considered only one mode of travel in the city.

The key literatures on the monocentric city models are summarized in Table 1. To our best knowledge, the dynamic congestion effect in a monocentric city with alternative transportation modes has not been studied previously. This paper addresses directly this problem.

Table 1. Developments in economic models of urban structure.

City model	Transportation modes	Congestion effect	Departure-time choice	Key references
Continuous	Highway	No	No	Alonso (1964); Mills (1967); Muth (1969)
Continuous	Highway	Yes	No	Strotz (1965); Solow (1972); Anas & Xu (1999); Li et al. (2013)
Core-Suburb	Highway	Yes	Yes	Arnott (1998)
Continuous	Highway	Yes	Yes	Gubins &Verhoef (2014); Takayama & Kuwahara (2016)
Continuous	Highway & Railway	No	No	Capozza (1973); Su & DeSalvo (2008)
Core-Suburb	Highway & Railway	Yes	No	Buyukeren & Hiramatsu (2016)
Core-Suburb	Highway & Railway	Yes	Yes	This paper

Thus the first contribution of this paper is to introduce dynamical departure-time choices into commuters travel behaviour in a monocentric city with two alternative transport modes, and to investigate the dynamic congestion effect on the economic activities of the city. We consider a discrete core-suburb city with a congested highway and a crowded railway. In this city, the residents' trip timing, travel mode, residential location and consumption choices all interact endogenously in order to maximize their utility. We analyse properties of the proposed city model, and investigate the effects of dynamic congestion and modal substitution on residential location choice and the city structure. Theoretical proofs of the results are provided.

Building on the above city model and the properties of the model, the second part of the paper is set out to address land-use and transport policy design problems. We show how a Pareto-improving policy solution can be achieved that improves the residents' utility whilst limiting urban sprawling.

Urban sprawl causes negative effects on a city: traffic congestion, pollution, eroding a city's open space and its surrounding green field, etc. Studies on road network patterns by Zhao et al. (2016) show a positive link between population distribution and land used for road network connection and length. A direct instrument to curb urban sprawl is to set a limit on the land use for housing. For example, London puts around it a greenbelt in the late 1930s to prevent city sprawl (Jun, 2004; Cohen, 1994), and similar practice has also been adopted in cities around the world such as in Moscow, New Delhi, Ottawa and Tianjin (Anas and Pines, 2008). Most of the US metropolitan areas use agricultural zoning to protect resource lands (Brabec and Smith, 2002). In a closed monocentric city with only one travel mode, studies by Pines and Sadka (1985), Wheaton (1998) and Brueckner (2007) showed that land use regulations are effective against urban externalities (e.g., traffic congestion) by reducing the city's spatial size and shifting the residents closer to the CBD. Buyukeren and Hiramatsu (2016) showed that the combined use of transit subsidy and UGB policies is effective for mitigating congestion effect in a city with a congested highway and crowding-free public transit.

The second contribution of this paper is a theoretical analysis on the effects of transit service improvement and land use policies in our proposed city model. We prove that the land use tax is an effective policy for anti-

congestion and anti-sprawl. However, the joint effects of transit improvement and land use policies on urban spatial structure and residents' utility are ambiguous; they depend on the trade-off between the centralizing effect caused by land-use tax and the decentralizing effect caused by the lower travel cost of the transit improvement policy.

Finally, looking from the prospects of both the city government for sustainable urban developments and the residents for utility maximization, we put forward a concept of Pareto-improving city development which improves the residents' utility without causing urban sprawling. We provide the conditions for the existence of Pareto-improving policies and present a method for deriving the policy solutions.

2. The model of a closed core-suburb city with congested highway and crowded railway

2.1. The model framework

We consider a closed monocentric city with two discrete residential locations: a core that contains a landless job centre and a suburb surrounded by green agriculture land (Fig. 1). Such a discrete city model is suitable to represent those areas with high-density population and homogeneous housing (such as in high-rising buildings). The core and the suburb are connected by a congested highway and a crowded railway. For simplicity without losing generality, the highway congestion is reflected by a capacity constrained bottleneck towards the core of the city, while the railway crowding is represented by body congestion in train carriages which causes discomfort to passengers. Passenger waiting time due to long train headway or failure to board due to low capacity are ignored here.
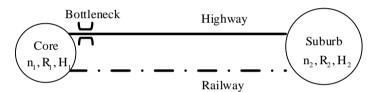

Fig. 1. A simple core-suburb city structure.

This simplified city model can be characterised as follows:

A1: The city's population is fixed, i.e. it is a 'closed' city. An exogenously determined N number of residents choose to live in one of the two locations. The residents can migrate freely between the core and the suburb. The number of residents in each area, denoted as n_1 and n_2 for the core and the suburb respectively, is determined endogenously, while $n_1 + n_2 = N$.

A2: The intra-zonal travel costs are ignored, e.g., the travel costs inside the core and the suburb are set to zero. Suburban residents incur travel costs as they travel between the suburb and work place in the core, they have a travel mode choice of: (a) transport mode: auto mode on the highway or transit mode using the railway, and (b) departure time to reduce congestion and schedule delay cost. The final travel cost is endogenously determined by their simultaneous mode choice and departure time decisions.

A3: All residents are assumed to be rational and homogenous with the same exogenously determined income level (Y). The income will be spent on transportation, housing and other consumption. The objective of the residents is to maximize their respective utilities within their budget constraint.

A4: The supply of land in the core (H_1) is assumed fixed, while the aggregate area of the suburb (H_2) is determined endogenously by the number of residents in the area and a unit size of housing (a lot size) (q_2). The latter is linked closely to the rent in the area. In the suburb, we assume that the residential lands are homogenous. In a competitive equilibrium situation, the residential land rent (R_2) is equal to that of agriculture (R_a), where R_a is an exogenous variable and represents the cost to convert one unit of agricultural land to one housing unit. The land is owned by absentee landlords and the land tax revenue is collected by government.

With the above defined city, we investigate the interplays between the residents' residential location and travel choices (mode and departure time), and the government's management policies. Fig. 2 illustrates some of the relationships. The residents' travel choices are influenced by the population density of the area as well as the government's policies on transit. Taking the endogenous travel cost as an input, the residential location equilibrium determines the size of the suburb, spatial population distribution, rental price in the core and the residents' utility level. The government sets the land use control and transit policies; both of which have direct impact on residents' location choice and travel choice, which in turn affect the urban spatial structure and utility level.
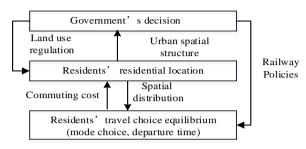

Fig. 2. The interactions in a core-suburb city.

Two significant behavioural equilibria emerge from the above interactions: (i) transportation equilibrium with mode and departure time choices; and (ii) urban spatial equilibrium with residential location choices. We present below analytically these equilibrium conditions in this core-suburb city model.

2.2. Equilibrium of mode and departure time choices

In this subsection, we analyze the commuting behaviours of suburban residents who travel between the suburb and the CBD. As assumed in A2, travel costs only be incurred between the core and suburb, while the cost inside the core and suburb are set to zero, for simplicity. The residents in the suburb have a combined choice of travel modes (using the highway or by transit) and departure time (affected by the highway bottleneck congestion and the crowding on the railway). We formulate below the travel costs for each mode first, before setting out the combined mode and departure-time equilibrium condition.

2.2.1. Travel cost using the highway

For the core-suburb city of Fig. 1, there is a bottleneck on the highway with capacity w at the entrance to the core. When the arrival rate exceeds this capacity, traffic queue forms. We apply the standard bottleneck model (Vickrey, 1969; Arnott et al., 1990) to derive the highway travel cost for departing at time t, which includes travel time cost, schedule delay costs and a monetary cost as follows:

$$C_h(t) = \alpha T_h(t) + \beta[t^* - T_h(t) - t]_+ + \gamma[t + T_h(t) - t^*]_+ + F_h \tag{1}$$

where $T_h(t)$ is the travel time at departure at time t which includes a constant free-flow travel time T_h and a time-dependent queuing time $T_w(t)$, t^* is the desired arrival time at the workplace, F_h represents a direct out-of-the-pocket cost (e.g. parking fee and petrol consumption). The parameters α, β and γ represents respectively the values of travel time, schedule delay early and schedule delay late, while a relationship $\beta < \alpha < \gamma$ holds according to Small (1982). x_+ is defined as $x_+ = \max\{x, 0\}$.

Let $[t_{hs}, t_{he}]$ be the highway commuters' departure time window, and define the peak period as $D_h = t_{he} - t_{hs}$. According to the standard bottleneck model, we can derive the departure rate of highway commuters as follows:

$$r_h(t) = \begin{cases} \alpha w/(\alpha - \beta), t \in \left[t_{hs}, t^* - T_h\right] \\ \alpha w/(\alpha + \gamma), t \in \left[t^* - T_h, t_{he}\right] \end{cases} \tag{2}$$

and the earliest and the latest departure times as:

$$t_{hs} = t^* - T_h - \frac{\gamma}{\beta + \gamma} \frac{n_{2h}}{w} \quad \text{and} \quad t_{he} = t^* - T_h + \frac{\beta}{\beta + \gamma} \frac{n_{2h}}{w}. \tag{3}$$

At equilibrium, all highway users should experience the same and minimal trip cost regardless of their departure times. This equilibrium highway travel cost is:

$$C_h = \alpha T_h + \delta n_{2h}/w + F_h. \tag{4}$$

where $\delta = \beta\gamma/(\beta + \gamma)$, and n_{2h} denotes the number of highway commuters from suburb.

The second term in Eq. (4) represents the queuing delay. It is a pure deadweight loss, and can be eliminated by a time-varying system-optimal toll (Gonzales and Daganzo, 2012; Amirgholy and Gonzales, 2016; Xiao et al., 2015a, 2015b). Note that substituting the queue delay cost with an optimal toll does not change the highway travel cost, thus will not change the mode choice and the equilibrium travel costs. The total revenue generated from the toll is $\kappa\delta(n_{2h})^2/w$, where κ represents an average number of days of travel per annum per person.

2.2.2. Travel cost using the railway

We assume that the rail travel times are reliable and constant between the suburb and the core. However, crowding and body congestion in train carriages affects transit users' departure time choices. The generalized travel cost of rail users covers four parts: free-flow travel time cost, schedule delay cost, in-carriage crowding cost (which is proportional to the number of residents in carriage and travel time), and the fare (including the normal fare part and subsidy part). The transit travel cost can be represented as follows:

$$C_r(t) = \alpha T_r + \beta\left(t^* - T_r - t\right)_+ + \gamma\left(t + T_r - t^*_r\right)_+ + \psi T_r \xi r_r(t) + (F_r - \phi), \tag{5}$$

where T_r is the free-flow travel time, $r_r(t)$ is the departure rate of rail commuters at time t, ξ is the train dispatching time interval, ψ is the in-carriage congestion parameter, F_r is the normal fare, and ϕ is railway subsidy. Other parameters or variables are the same as those in Eq. (1). The number of passengers who board on a train departing at time t_0 is $\int_{t_0 - \xi}^{t_0} r_r(t) dt$. For simplicity, we follow the assumption in Wu and Huang (2014) that the departure rate of rail commuters is continuous and constant during one train dispatching interval, thus we have $\int_{t_0 - \xi}^{t_0} r_r(t) dt \approx \xi r_r(t_0)$. The crowding cost of a train is directly proportional to the number of passengers boarded that train and the running time of the train.

Let $\left[t_{rs}, t_{re}\right]$ be the rail commuters' departure time window, and the peak period on the railway be $D_r = t_{re} - t_{rs}$. According to user equilibrium condition, the departure rate of rail commuters can be derived as:

$$r_r(t) = \begin{cases} \beta(t - t_{rs})/(\psi\xi T_r), t \in \left[t_{rs}, t^* - T_r\right] \\ \gamma(t_{re} - t)/(\psi\xi T_r), t \in \left[t^* - T_r, t_{re}\right] \end{cases} \tag{6}$$

and the earliest and the latest departure times as:

$$t_{rs} = t^* - T_r - \sqrt{2\delta\psi\xi T_r n_{2r}}/\beta \quad \text{and} \quad t_{re} = t^* - T_r + \sqrt{2\delta\psi\xi T_r n_{2r}}/\gamma. \tag{7}$$

where $n_{2r} = n_2 - n_{2h}$ is the number of residents travelling by the railway.

At departure-time equilibrium, the generalized travel cost of a transit commuter is:

$$C_r = \alpha T_r + \sqrt{2\delta\psi\xi T_r n_{2r}} + (F_r - \phi). \tag{8}$$

2.2.3. Equilibrium condition of travel mode choice

At equilibrium, all residents who travel from the suburb to the core should have the same and minimal trip cost, regardless of their transport modes, i.e.,

$$C_h = C_r \tag{9}$$

To ensure both modes are used, the mode-choice equilibrium requires the following two conditions:

$$C_h(n_2) > C_r(0) \text{ and } C_h(0) < C_r(n_2) \tag{10}$$

Eq. (10) states that if all suburb residents choose to travel by only one mode, their travel cost would be higher than the cost of unused mode.

2.3. General urban spatial equilibrium

According to assumption A3, each resident chooses the consumption of housing and non-housing composite good to maximize his/her own utility under budget constraint. The direct utility function of the commuter living in area i is given by $U(z_i, q_i)$, where z_i is the annual consumption of non-housing composite good and q_i is the unit residential housing lot rented by the commuters at an annual cost of R_i per unit. $U(z_i, q_i)$ is twice-differentiable, strictly increasing and strictly concave in z_i and q_i. The utility maximization problem can be formulated as:

$$V(R_i, Y - TC_i) = \max U(z_i, q_i)$$

subject to

$$z_i + R_i q_i = Y - TC_i \tag{11}$$

where $V(\cdot)$ is the indirect utility function, the price of non-housing good is normalized to \$1, TC_i is the annual generalized travel cost in area i while $TC_1 = 0$ according to A1, and $TC_2 = TC_h = 2\kappa C_h$ or $TC_2 = TC_r = 2\kappa C_r$ with the multiplier 2 denoting a daily round-trip and κ the average annual number of trips between the residential location and the CBD per resident. From Eq. (11), $q_i = (Y - z_i - TC_i)/R_i$, we know that $\partial q_i / \partial R_i < 0$.

The compensated demand functions of housing and non-housing good can be written as $q(R_i, u)$ and $z(R_i, u)$, where u is the endogenously determined equilibrium utility level. For simplicity, we use $E(R_i, u)$ to represent the minimum expenditure function, i.e., $E(R_i, u) = z(R_i, u) + R_i q(R_i, u)$.

Similar to Arnott (1998), the closed-city model at equilibrium has the following characteristics.

(a) Migration equilibrium:

$$V(R_1, Y - TC_1) = V(R_2, Y - TC_2) = u \tag{12}$$

(b) Land-market equilibrium conditions:

$$n_1 q(R_1, u) = H_1 \tag{13}$$
$$n_2 q(R_2, u) = H_2 \tag{14}$$

(c) Budget constraints:

$$E(R_1, u) - Y = 0 \tag{15}$$

$$E\left(R_2, u\right) + TC_h - Y = 0 \tag{16}$$

$$E\left(R_2, u\right) + TC_r - Y = 0 \tag{17}$$

(d) Population conservation:

$$n_1 + n_2 = N. \tag{18}$$

where the suburban land rent per unit is $R_2 = R_a$. Eq. (12) states that residents living in the core and the suburb should have the same utility at equilibrium. Eqs. (13) and (14) are the land market clearing conditions, which give the endogenously determined number of residents and the lot size in each area and the area of the suburb. Eqs. (15) – (17) give the budget constraints of the core residents and the suburban residents who use highway and railway, respectively. The annual total travel costs, considering a round trip per travelled day, are:

$$TC_h = TC_r = 2\kappa C_h = 2\kappa C_r. \tag{19}$$

Remark. In our paper, residents maximize their utility by choosing the optimum residential location, lot size and amount of non-housing goods within the budget constraint. We do not consider the influence of environmental amenities on people's choices. In reality, however, different residents' preferences to the density and the quality of housing would vary. Therefore, environmental externality would have a significant influence on residential location choice behaviours. Considering the effect of location-dependent environmental amenities, studies have shown that the city structure and the spatial population distribution depend on the trade-off between traffic congestion and environmental externalities (e.g., Verhoef and Nijkamp, 2008; Chen et al., 2016). Incorporating the effect of environmental externalities into the urban model is clearly very important and warrants further research.

3. General properties

Proposition 1. For a given rail fare, a decrease of transit dispatching headway leads to: (a) a decrease of the core population; (b) a decrease of the number of highway commuters; (c) an increase of the number of rail commuters; (d) a decrease of the core land rent; (e) an increase of the equilibrium utility level; (f) an increase of suburb land area, (g) a decrease of the highway peak-period, and (h) a decrease of railway travel peak-period.

Proof: See Appendix A.

The economic interpretation of Proposition 1 is that in our core-suburb city model with alternative and congested transport modes, increasing service frequency provides more capacity and reduces crowding, which in turn attracts more highway commuters to switch to rail. Then the suburban residents will have more disposable income due to the reduced travel cost, which also attracts more core residents to relocate in the suburb. Then, according to Eq. (17) and the Shepherd's Lemma $\partial E_1 / \partial R_1 = q_1$, the core land rent will decrease. With a fixed core land size (assumption A4) and according to Eq. (14), the core resident can now consume a larger lot size. As the utility is a strictly increasing function of lot size, the equilibrium utility will be higher. Following Eq. (3) and with Proposition 1(b), the highway peak period is reduced. Proposition 1(h) suggests that the combined effect of a reduced transit headway and increased rail users is to reduce the rail peak period.

Proposition 2. Under a given transit headway, an increase of transit subsidy leads to the following: (a) a decrease of core population; (b) a decrease of the number of highway commuters; (c) an increase of the number of rail commuters; (d) a decrease of core land rent; (e) an increase of the equilibrium utility level; (f) an increase of suburb land area; (g) a decrease of peak-period on highway; and (h) an increase of peak-period on railway.

Proof: See Appendix B.

Proposition 2 suggests that increasing transit subsidy yields similar results to that of reducing service headway except for its effect on rail peak period which, according to Eq. (8), would increase with increasing suburb rail commuters (Proposition 2(c)).

We list in Table 2 comparative statistics of the above results. It shows clearly that, with both transit improvement methods, the combined modal substitution and dynamic congestion effects on the city are to increase the utility level of the residents, decrease the core population, and increase the suburb lane use causing urban sprawl. This result is different to that based on a static city model with mode choice. In the congested city models with modal choices but without dynamical departure-time choice, it found that the effects of congestion toll or public transport improvements on city spatial structure are ambiguous.

Table 2. Comparative statistics for improving public transport provisions

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
	n_1	n_{2h}	n_{2r}	R_1	u	H_2	D_h	D_r
Decreasing ξ	-	-	+	-	+	+	-	-
Increasing ϕ	-	-	+	-	+	+	-	+

Table 2 also shows that the two transit improvement methods have one different impact: that is on the peak transit travel period. As discussed above, the combined effect of a reduced transit headway and increased rail users is to reduce the rail peak period, whilst the singular effect of increased rail users as a result of rail subsidy is to increase the rail peak period. Fig. 3 illustrates the effects of transit improvements on the temporal distribution of railway travel demand, with an exogenously given suburb travel demand and an endogenously determined demand. It shows that, when the whole city system is considered and transit demand is endogenously determined with high dispatching headway or low subsidy of transit, the transit improvement policies have a more significant impact on commuters' travel behaviour than if transport is considered in isolation.
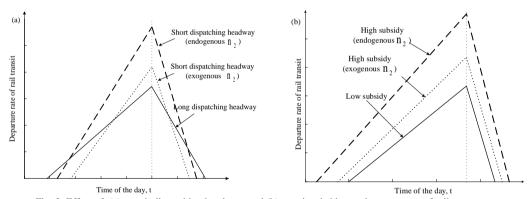

Fig. 3. Effect of: (a) transit dispatching headway, and (b) transit subsidy, on departure rate of railway commuters.

4. Pareto-improving land-use and transit policies

As we can see from Section 3 that improving transit services (with shorter service headway and/or subsidy) increases residents' equilibrium utility, but at a cost of increased suburb land use and urban sprawl. In this section, we analyse the trade-off between the two effects and examine policy designs that encourage urban economic activities whilst improving the residents' utility.

4.1. Land-use and transport policy instruments

We consider four planning regimes that are based on the land-use and transit instruments available to local government, and explore their effects on our city. The four regimes are defined as follows:

(a) Land use tax $\{s\}$: the local government imposes a tax $s > 0$ on the suburb land rent. Then, the suburban land rent per unit becomes $R_2 = R_a + s$.

(b) Cross land use taxation and rail service enhancement $\{s, \xi^-\}$: the government sets the land-use tax $s > 0$ while simultaneously reduces the transit service headway ξ.

(c) Cross land use taxation and direct railway fare subsidy $\{s, \sigma\}$: the government sets the land-use tax $s > 0$ and redistributes a proportion of the tax revenue equally among rail users through a transit subsidy. The subsidy ϕ can be designed as $\phi = \sigma n_2 q_2 s / (2\kappa n_{2r})$, where $0 < \sigma < 1$ is the proportion of land use tax used for transit subsidy.

(d) Cross optimal road toll and direct railway fare subsidy $\{\tau, \sigma\}$: the government sets a time-varying toll τ to eliminate queuing, then redistributes a proportion of the toll revenue equally among rail users through a transit subsidy. The subsidy can be designed as $\phi = \sigma \delta (n_{2h})^2 / (2w n_{2r})$.

Imposing a land tax to the suburb is to make it more expensive to live, which in turn would make the core more attractive and lead to a centralizing effect. With a fixed core land size H_1, the increased core population would reduce the lot size q_1. Since the utility function is strictly increasing with lot size, the core equilibrium utility would then reduce. This behaviour under land-use instrument $\{s\}$ can be summarised in the following Propositions.

Proposition 3. The effect land-use tax on urban spatial structure is to centralize the core city, and decreases the equilibrium utility of the residents.

We have shown in Proposition 1 that a reduction in transit headway would result in an increase in equilibrium utility and an increase in urban sprawl. Thus, with policy $\{s, \xi^-\}$, when a land-use tax is combined with a reduction of transit headway, the results on urban structure and economics become ambiguous.

Similarly, with policy $\{s, \sigma\}$, there is a trade-off between the centralizing effect caused by land-use tax (as in Proposition 3) and the decentralizing effect caused by transit subsidy (Proposition 2). However, in this case, the level of subsidy (ϕ) is an endogenous variable linked to the level of land-use tax. Nonetheless, we record in Proposition 4 below, that the combined policy effects on the city are still ambiguous.

Proposition 4. In the above core-suburb congested dual-mode city, using land-use tax to subsidy rail transit, the effects on urban spatial structure are ambiguous.

Proof: According to Eq. (14), we get

$$\Delta H_2 = \Delta \big(n_2 q (R_a, u) \big) = q_2 \Delta n_2 + n_2 q_{2u} \Delta u + n_2 \frac{\partial q_2}{\partial R_2} \Delta R_2$$

where $\partial q_2 / \partial R_2 < 0$ and $\Delta R_2 > 0$ with the addition of the land-use tax. We consider two scenarios below:

(i) $n_1^{s\sigma} \geq n_1$. Then $n_2^{s\sigma} \leq n_2$, i.e., $\Delta n_2 \leq 0$. According to Eq. (13), we get $q_1^{s\sigma} \leq q_1$. Then we get $\Delta u \leq 0$ since the utility function increases with lot size. Thus, $\Delta H_2 \leq 0$.

(ii) $n_1^{s\sigma} < n_1$. Then we get $\Delta n_2 > 0$. We also can easily derive that $u^{s\sigma} > u$, i.e., $\Delta u > 0$. According to the equation of ΔH_2, the sign of ΔH_2 is ambiguous. □

Imposing a system-optimum toll does not alter residents' travel mode choice, hence will not change the city structure if the toll revenue is not redistributed. This result is different to that based on a static city model with mode choice, which shows that the effect of congestion toll on city spatial structure is ambiguous (Buyukeren and Hiramatsu, 2016). Proposition 5 shows the effects of redistribution of toll revenue to subsidy transit on urban spatial structure and residents' utility level.

Proposition 5. In the above core-suburb congested dual-mode city, using toll revenue to subsidy rail transit, causes increase of the equilibrium utility level and increase of suburb land area.

Proof: Using the Reductio ad absurdum method, we assume that $n_1^{r\sigma} \geq n_1$, then $n_2^{r\sigma} \leq n_2$. According to Eq. (13), we get $q_l^{r\sigma} \leq q_l$. Then we get $u^{r\sigma} \leq u$. Then, $TC_h^{r\sigma} \geq TC_h$ and $TC_r^{r\sigma} \geq TC_r$ can be obtained according to Eq. (16) and (17), respectively. That is, $n_{2h}^{r\sigma} \geq n_{2h}$ and $n_{2r}^{r\sigma} \geq n_{2r}$, which are not consistent with the assumption $n_2^{r\sigma} \leq n_2$. $\square$

Table 3 summarises the effects of four policy instruments on urban structure and residents' utility level.

Table 3. Summary of the effects of land-use and transport policy instruments.

	$\{s\}$	$\{s, \xi^-\}$	$\{s, \sigma\}$	$\{\tau, \sigma\}$
u	-	?	?	+
H_2	-	?	?	+

Clearly, there is a trade-off between the centralizing effect caused by land-use tax and the decentralizing effect caused by improvements of transit services. If the centralizing effect dominates over the suburbanization effect, the residents' utility level decreases and the aggregate land use of the suburb reduces. Otherwise, the residents' utility level increases and the suburb area sprawls out.

4.2. Definition and condition of a Pareto-improving policy solution

We consider a government who aims at simultaneously mitigating congestion, maintaining a sustainable city growth and making its residents better off by planning land-use and transport policy instruments. We define a policy that achieves such objectives as Pareto-improving.

Definition 1. A design scheme is said to be Pareto-improving if it holds that

$$u^p \geq u^e \quad \text{and} \quad H_2^p \leq H_2^e. \tag{20}$$

where u^p and H_2^p, u^e and H_2^e are the utility level and size of suburb land area corresponding to Pareto-improving equilibria and the base equilibrium case. The left part of condition (20) means that, with the Pareto-improving policy, the utility level of each resident is improved, i.e., everyone is made better off. Meanwhile the right part of condition (20) states that the size of suburban land area is not expanded, i.e. no urban sprawl.

Lemma 1. If Pareto-improving transit scheme exists, the number of the core residents under the scheme decreases, namely $n_1^p < n_1^e$.

Proof: See Appendix C.

4.3. Existence of Pareto-improving designs

From Table 3, it is clear that land use tax $\{s\}$ and policy $\{\tau, \sigma\}$ have a definite effect on urban spatial structure and residents' utility level, respectively. Therefore, Pareto-improving policy solutions cannot exist under the policy (a) and policy (d). However, we show in Table 3 the ambiguous effects of policy instruments $\{s, \xi^-\}$ and $\{s, \sigma\}$ on our city structure. In this section, we provide conditions for the existence of Pareto-improving designs under these two instruments.

(i) Conditions for Pareto-improving transit headway - policy (b)

According to Proposition 3, under a given land use tax s, if there exists a headway that keeps the utility level, then the suburban land area shrinks. When the utility level keeps unchanged, the number of the core residents and

suburb residents are also unchanged, but the land consumption in the suburb decreases due to high land rent, then the suburb size shrinks. From Proposition 1, we know that the utility level improves and the aggregate suburb land use increases as headway decreases. Therefore, the range of the Pareto-improving transit headway must follow:

$$\xi^p \in \left[\underline{\xi}, \overline{\xi}\right] \tag{21}$$

where $\underline{\xi}$ represents the shortest dispatching headway and it satisfies $H_2\left(s, \underline{\xi}\right) = H_2\left(s = 0, \xi_0\right)$, where ξ_0 is the base headway, while $\overline{\xi}$ is the longest dispatching headway and satisfies $u\left(s, \overline{\xi}\right) = u\left(s = 0, \xi_0\right)$.

Lemma 2. For any given s, $\underline{\xi}$ and $\overline{\xi}$ if exist, are unique.

Proof: See Appendix D.

Proposition 6. With an increase of land use tax s, the upper and lower boundaries of Pareto-improving transit headway ($\overline{\xi}$ and $\underline{\xi}$) both decrease.

Proof: See Appendix E.

Proposition 6 implies that an increase in land use tax must be accompanied by a re-investment in reducing transit headway to achieve Pareto-improving solutions. In another word, since a reduced transit headway leads to city expansion (Proposition 1), the government would need to set a higher land use tax to control the urban sprawl.

(ii) Conditions for Pareto-improving transit subsidy ratio - policy (c)

In policy instrument $\{s, \sigma\}$, the government redistributes a proportion of the tax revenue equally among rail users through a transit subsidy. First, we examine the effect of σ under a given s.

Proposition 7. In the above core-suburb city, given a land use tax, using tax revenue to subsidize railway commuters leads to the following: an increase of subsidy ratio σ leads to increase in the equilibrium utility level and an increase in suburb land area.

Proof: See Appendix F.

Proposition 7 implies that the higher transit subsidy, the more disposable income the railway commuters get, which in turn attracts more highway commuters to switch to railway and more core residents to relocate to the suburb. Thus, the bigger lot size the remaining core residents consume according to Eq. (13), the higher utility there is. In the suburb, each resident also consumes a larger lot size because utility is higher, then the size of the suburb becomes bigger.

According to Proposition7, therefore, the range for the Pareto-improving transit subsidy ratio must follow:

$$\sigma^p \in \left[\underline{\sigma}, \overline{\sigma}\right] \tag{22}$$

where the smallest $\underline{\sigma}$ satisfies $u\left(s, \underline{\sigma}\right) = u\left(s = 0\right)$, the largest subsidy ratio $\overline{\sigma}$ satisfies $H_2\left(s, \overline{\sigma}\right) = H_2\left(s = 0\right)$ if $H_2\left(s, 1\right) > H_2\left(s = 0\right)$, otherwise, $\overline{\sigma} = 1$.

Lemma 3. For any given s, $\underline{\sigma}$ and $\overline{\sigma}$ if exist, are unique.

Proof: See Appendix D.

4.4. Procedure for calculating the Pareto-improving solution

For a given tax s, the Pareto-improving solutions can be derived by firstly solve for the values $\underline{\xi}$ and $\bar{\sigma}$, and from which to solve for values $\bar{\xi}$ and $\underline{\sigma}$. The procedure is described below.

Algorithm 1: Procedure for solving Pareto-improving transit headways and subsidies

---

Set the initial equilibrium solutions H_2^e and u^e

Solve for the value of $\underline{\xi}$ and $\bar{\sigma}$

   Derive values of n_1^p, n_2^p, u^p and R_1^p by solving Eqs. (13), (14), (15) and (18)

   Solve n_{2h}^p and n_{2r}^p from Eq. (16) and population conservation condition $n_{2h}^p + n_{2r}^p = n_2^p$

   Obtain $\underline{\xi}$ by solving Eq. (19)

   Solve $q(R_2, u)^p$ according to Eq. (14)

   Obtain $\bar{\sigma}$ by solving Eq. (19)

Solve for the value of $\bar{\xi}$ and $\underline{\sigma}$

   Derive values of n_1^p and n_{2h}^p by solving Eq. (13) and $E(R_2, u)^p + TC_h^p = E(R_2, u)^e + TC_h^e$

   Solve n_{2r}^p from population conservation condition $n_{2h}^p + n_{2r}^p = n_2^p$

   Obtain $\bar{\xi}$ by solving Eq. (19)

   Solve $q(R_2, u)^p$ according to Eq. (16) or (17)

   Obtain $\underline{\sigma}$ by solving Eq. (19)

---

5. Model applications

In this section, we present numerical results to illustrate the effect of urban land use taxation and transportation policies on the population choices and the economic performance of our core-suburb city. To allow for numerical calculation, we adopt the Cobb–Douglas form of household direct utility function:

$$U(z_i, q_i) = z_i^{\theta} q_i^{(1-\theta)}, \qquad \theta > 0 \tag{23}$$

Following the spatial equilibrium conditions of Section 2.3, the unit price of a housing lot R_i and the lot size q_i as functions of the equilibrium utility u can be derived as following:

$$R_i(u) = \theta^{\frac{\theta}{1-\theta}}(1-\theta)(Y - TC_i)^{\frac{1}{1-\theta}} u^{-\frac{1}{1-\theta}}; \quad q_i(u) = \theta^{-\frac{\theta}{1-\theta}}(Y - TC_i)^{-\frac{\theta}{1-\theta}} u^{\frac{1}{1-\theta}} \tag{24}$$

The reference parameter values used in our numerical analysis are given in Table 4.

Table 4. The reference parameter values.

Symbol	Value	Symbol	Value	Symbol	Value
N	50000residents	Y	$60000	R_a	30000$/acre
H_1	800acres	θ	0.8	t^*	9:00am
α	1.2$/min	β	0.8$/min	γ	3$/min
T_h	20min	F_h	$12	w	220veh/min

T_r	35min	F_r	$6	ψ	0.0001$/(min*pax)
ξ	5min	κ	250		

5.1. Results of the joint land use tax and transit headway instrument

We first illustrate the effects of transit headway (ξ) on mode choice pattern, peak period of travel, population spatial distribution, core land rent, aggregate land use of the suburb utility level of each resident. A range of transit headway values, from 1, 2, …, 5 min, are tested and the results are presented in Table 5.

Table 5. The effect of the rail headway (ξ).

ξ (min)	5	4	3	2	1
n_1	41561	41195	40639	39688	37658
n_{2h}	6475	6435	6373	6264	6021
$[t_{hs},t_{he}]$ (min)	[-23.24,6.20]	[-23.09,6.16]	[-22.87,6.10]	[-22.48,5.99]	[-21.61,5.76]
n_{2r}	1964	2370	2988	4048	6321
$[t_{rs},t_{re}]$ (min)	[-8.24,2.20]	[-8.09,2.16]	[-7.87,2.10]	[-7.48,1.99]	[-6.61,1.76]
TC ($)	27295	27237	27148	26991	26643
R_1 ($/acre)	623412	617932	609591	595315	564870
u	2522.75	2527.21	2534.09	2546.13	2573.00
H_2 (acres)	1840.06	1923.11	2050.13	2269.31	2744.60

Note: In this numerical example, $s = 0$.

It can be seen that with a shorter headway, more suburb residents choose to use the railway at lower travel cost. They then have more disposable income which they can use to consume more land and their utility level increases. Thus, with reducing transit headway, the decentralizing effect becomes stronger. As shown in Table 5, the core population decreased from 41561 to 37658 when ξ is reduced from 5 min to 1 min. This also reduces the core land rent from 623412 $/acre to 564870 $/acre. The relocation of core residents to the suburb then increases utility level from 2522.75 to 2573.00. In addition, we can see that shorter headway reduces the number of highway commuters from 6475 to 6021, which reduces the peak period on the highway. While the number of railway users is increased from 1964 to 6321, the peak period on railway is still reduced.

Next we examine the effect of a reduced railway headway combined with a land use tax on utility level and the city spatial structure. Fig. 4 shows the changes in equilibrium utility and the suburb land size with respect to land-use tax and transit headway. It can be seen from Fig. 4(a) that, for a given land-use tax, the equilibrium utility increases with decreasing dispatching time interval. In order to achieve the same utility level, the higher the land-use tax is, the shorter the dispatching time interval is required. The dashed curve defines the utility level ($u = 2522.75$) with the reference headway $\xi = 5$ min when land-use tax and railway subsidy are not available. At this utility level, if a land tax of $s = 1000$ $/acre is introduced, then an upper bound of transit headway of $\bar{\xi} = 4.05$ min is required; while if $s = 2000$ $/acre, the upper headway is reduced to $\bar{\xi} = 3.30$ min, implying more frequent transit is required.

Fig. 4(b) is a contour plot of the suburb size with a range of combinations of land use tax and railway headway. For a given land-use tax, as headway decreases, the suburb size increases, which is consistent with Proposition 2. We also find that under a given suburb size H_2, the railway headway decreases with land use tax. For example, for $H_2 = 1840.06$ acres, if $s = 1000$ $/acre, then $\underline{\xi} = 3.56$ min; while if $s = 2000$ $/acre, $\underline{\xi} = 2.60$ min.
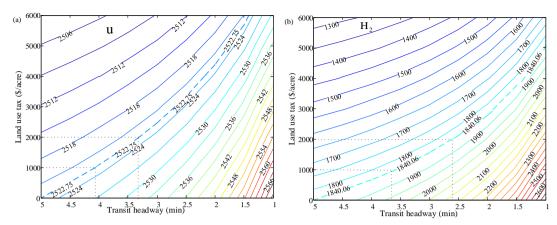
Fig.4. Change of: (a) equilibrium utility, and (b) suburban land size, with $\{s, \xi\}$.

Table 6 lists the values of Pareto-improving headways under different land use tax s. The Pareto-improving headway solution lies within $[3.56, 4.05]$ min when land use tax is $s = 1000$ \$/acre, while with a land tax of $s = 2000$ \$/acre, the upper and the lower boundaries both decrease to be $[2.60, 3.30]$ min, which is consistent with Proposition 5. This shows that a higher land use tax must be accompanied by a lower transit headway.

Table 6. Pareto-improving headway solutions with respect to different land use tax s.

s (\$/acre)	1000	2000	3000	4000	5000	6000
$\underline{\xi}$ (min)	3.56	2.60	1.90	1.40	1.03	0.75
$\overline{\xi}$ (min)	4.05	3.30	2.70	2.20	1.80	1.39

Note: In this numerical example, $H_2^e = 1840.06$ acres and $u^e = 2522.75$.

5.2. Results of the joint land use tax and rail transit subsidy instrument

In this subsection, we first examine, under a given level of land tax, the effect of subsidy ratio on the urban structure. Table 7 summarizes the results of mode choice pattern, population spatial distribution, the core land rent, the aggregate land use of the suburb and utility level of each resident, with the subsidy ratios range from 0 to 1.0 at interval of 0.2. We can see that when subsidy ratio σ increases, the suburbanization effect become stronger and the centralization effect become weaker. At $\sigma = 0.2$, the core population is 41708 compared with 41293 at $\sigma = 1$. While at $\sigma = 0.8$ the core population is 41346 compared with 41293 at $\sigma = 1$. When σ is small, the centralization effect caused by land use tax dominates the suburbanization effect caused by redistributed tax revenues to railway users. At $\sigma = 0.2$, the number of core residents increases from 41561 to 41708, and the suburban residents decreases from 8439 to 8292. The suburbanization effect increases as σ increases. At $\sigma = 0.6$, the core residents decreases and the suburban residents increases. What's more, the effect of railway subsidy on the suburban land area is ambiguous. The suburban land area shrinks 1840.06 acres to 1759.89 acres at $\sigma = 0.2$ while it expands from 1840.06 acres to 1851.79 acres at $\sigma = 1$.

Table 7. The effect of the subsidy ratio (σ).

σ	0	0.2	0.4	0.6	0.8	1
n_1	41870	41708	41560	41423	41346	41293
n_{2h}	6359	6341	6325	6310	6301	6296
$[t_{hs}, t_{he}]$ (min)	[-22.82,6.09]	[-22.76,6.07]	[-22.70,6.05]	[-22.64,6.04]	[-22.59,6.02]	[-22.54,6.01]
n_{2r}	1771	1950	2114	2267	2353	2412

$[t_{rs}, t_{re}]$ (min)	[-7.82,2.09]	[-8.21,2.19]	[-8.55,2.28]	[-8.85,2.36]	[-9.13,2.43]	[-9.38,2.50]
n_2	8130	8292	8440	8577	8654	8707
TC ($)	27128	27102	27079	27057	27045	27037
R_1 ($/acre)	628051	625621	623405	621339	620186	619388
H_2 (acres)	1724.16	1759.89	1792.52	1822.97	1840.00	1851.79
u	2519.01	2520.97	2522.76	2524.43	2525.37	2526.02

Note: In this numerical example, $s = 1000$ $/acre.

Fig. 5 present contour plots of the resident's equilibrium utility level and the suburb land size with varying combinations of land-use taxes and subsidy ratios. It can be seen (in Fig. 5(a)) that for a given land-use tax, as the transit subsidy ratio increases, the resident's equilibrium utility level increases, which is consistent with Proposition 6. We also find that, under a given utility level, the subsidy ratio does not always increase with land use tax. From Fig. 5(b), we can see that under a given land use tax, the aggregate land use of the suburb increases with the subsidy ratio, which is consistent with Proposition 6. However, the change trend of H_2 with s is ambiguous under the same subsidy ratio. Because the subsidy railway commuters obtained not only depends on total land use tax revenue but also depends on the number of railway commuters.

It is worth noting also from Fig 5(a) that the relationships between tax and equilibrium utility, for a given subsidy ratio, are not straight-forward. At lower subsidy ratios (below 0.35), the utility levels decrease with tax. The relationship is more complex at higher subsidy ratios.

Fig. 6 plots the utility levels vs land-use tax for a given subsidy ratio. We can see that, at higher subsidy ratios (at $\sigma = 0.4$ and above), the utility level - land tax forms a concave curve. This implies that, for high subsidy ratios, there exits an optimal land-use tax at which the residents' utility level is at its maximum. The optimal tax and the associated maximum utility are also shown in Fig. 6. It can be seen that the optimal tax and the associated maximum utility increase with the subsidy ratio. When the subsidy ratio $\sigma = 0.35$, the maximum utility is obtained with $s = 0$. Therefore, when subsidy ratio is less than 0.35, the utility levels decrease with tax. In addition, according to Proposition 6, the utility level increases with the subsidy ratio under a fixed land use tax. Therefore, there exists a land use tax and using all tax revenue to subsidy transit makes residents' utility level maximize. For this example city, that optimal land tax is 7400 $/acre.
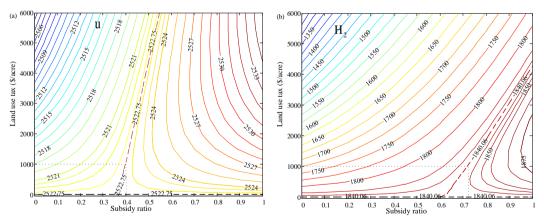

Fig.5. Change of: (a) utility level, and (b) suburban land area, with the land-use tax and the subsidy ratio.
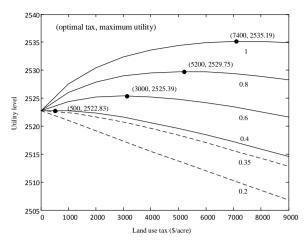
Fig.6. Changes of utility level with land-use tax under different subsidy ratios.

Table 8 gives the values of Pareto-improving subsidy ratio with different land use tax s. It shows that if the government sets a higher land use tax, it must also give more to subsidy railway commuters in order not to reduce the base utility level $u^e = 2522.75$, which is consistent with the dashed curve in Fig.5 (a). In addition, we find if the government set a high land use tax, it is also efficient to curb urban sprawl when all tax revenue to subsidy railway commuters.

Table 8. Pareto-improving subsidy ratio solutions with respect to different land use tax s.

s ($/acre)	1000	2000	3000	4000	5000	6000
$\underline{\sigma}$	40.0%	43.4%	46.0%	49.2%	51.8%	54.8%
$\bar{\sigma}$	71.7%	80.9%	89.6%	97.8%	100%	100%

Note: In this numerical example, $H_2^e = 1840.06$ acres and $u^e = 2522.75$.

6. Summary

In this paper, we introduce dynamical departure-time choices into commuters' travel behaviour in a monocentric city with two alternative transport modes. We analyse the dynamic congestion effect on the economic activities of the city where the residents maximize their utility under budget constraint. Their choices on household consumption, residential location, and transport are the results of endogenous optimized behaviour. In travel equilibrium, they choose between competing travel modes, and time of departure, to minimise congestion costs. Giving the endogenous equilibrium travel cost, land market equilibrium and migration equilibrium together determine the urban spatial structure (which includes the housing lot size, land rent, and spatial population distributions) and the maximised equilibrium utility.

This is a more realistic city setup than those reported in existing monocentric city literature. We formulate analytically our proposed city model, derive and provide theoretical proofs of the properties of the model. We show that the improvement of transit service has a definitive effect on the city structure: that it limits the centralization effect, attracts more residents to relocate in the suburb, thus leads to urban sprawl. This is in contrast to results based on city models with only highway travel mode: there mitigating congestion externality is found to have a centralising effect, i.e. a densification of the population towards the city centre. Our city model result is also different to that based on models which only consider travel mode choice, but not dynamic congestion effect on departure-time choices; the latter shows the effects of congestion externality on the city structure and population distribution are ambiguous.

Building on our city model and the derived properties, we explore transit policy designs in combination with land

use tax to rationalize the urban spatial size and promote residents' utility level. It is found that if the government improve development of transit at a high level, they must either accept a high level of urban sprawl or implement strict land use regulation to curb urban sprawl. We show that it is possible to design a Pareto-improving land-use and transit improvement policy which benefits the residents without causing urban sprawl. We provide theoretically the conditions for the existence of such policies and a practical solution method to derive the policy settings.

Although it has been shown that the proposed model has well-defined properties and provides useful insights for the sustainable development of the city, some important features of modern cities have been omitted in our model and should be considered in the future studies.

Firstly, for analytical tractability, our proposed city configuration is highly simplified to a two-zone monocentric city. This discrete core-suburb city model is suitable to characterize a city with two highly concentrated areas (i.e. high residential density) and homogenous lot size. However, in many modern cities, the residential locations are continuously distributed and there might be more than one business centre and residential location areas. Therefore, it would be worthwhile to adopt a continuous city model or polycentric city structure to study the relationship between land use and transportation.

Secondly, in this paper, we only consider the commuting costs for travelling between the suburb and the core; the travel costs within the suburb are ignored. This assumption is applicable under certain situations such as where the travel costs are homogeneous for all residents due to high-density population. In reality, however, the expansion of the city would increase the city size, then the gap between travel distances of residents resided in different locations within the suburb become bigger, which could have a significant effect on residents' travel behaviours. To enhance the realism of the model, intra-zonal travel behaviours warrants further research.

Thirdly, the environmental externality is ignored in this paper. In fact, the environmental externality can affect households' residential location choices and city spatial structure (Verhoef and Nijkamp, 2008; Li and Peng, 2016). Ignoring it may lead to underestimate the externality level that residents actually cause, which then lead to a distorted urban structure. Further studies can be carried out to incorporate the influence of environmental externality (e.g., negative effects incurred by residential density and vehicle emissions) so as to create a sustainable urban city system and correct the distortion.

Acknowledgements

Appendices

In the following proofs, we use abbreviated notion as follows: $E_{iu} = \partial E_i / \partial u$, $E_{R_i} = \partial E_i / \partial R_i = q_i$, $q_{iu} = \partial q_i / \partial u$, $q_{R_i} = \partial q_i / \partial R_i$, $A_1 = 2\kappa\delta / w > 0$, $A_2 = \kappa\sqrt{2\delta\psi\xi T_r / n_{2r}} > 0$ and $A_3 = \kappa\sqrt{2\delta\psi T_r n_{2r} / \xi} > 0$.

Appendix A: Proof of Proposition 1

Proof: Setting $\phi = 0$, Eqs. (15), (16), (19) and (13) totally differentiated with respect to ξ , getting

$$
\begin{bmatrix}
0 & q_1 & E_{1u} & 0 \\
0 & 0 & E_{2u} & A_1 \\
A_2 & 0 & 0 & A_1 + A_2 \\
q_1 & n_1 q_{1R_1} & n_1 q_{1u} & 0
\end{bmatrix}
\begin{bmatrix}
dn_1 / d\xi \\
dR_1 / d\xi \\
du / d\xi \\
dn_{2h} / d\xi
\end{bmatrix}
=
\begin{bmatrix}
0 \\
0 \\
A_3 \\
0
\end{bmatrix}
\tag{A1}
$$

The determinant of (A1) is negative: $\det_\xi = E_{1u} A_1 A_2 n_1 q_{R_1} - E_{2u} (A_1 + A_2) q_1^2 - A_1 A_2 q_{1u} H_1 < 0$. Then using

Cramer's rule, we find:

(a) $\dfrac{dn_l}{d\xi} = \dfrac{A_l A_3 n_l \left(E_{lu} q_{R_l} - q_{lu} q_l \right)}{\det_\xi} > 0$; (b) $\dfrac{dn_{2h}}{d\xi} = \dfrac{-E_{2u} A_3 q_l^2}{\det_\xi} > 0$; (c) $\dfrac{dn_{2r}}{d\xi} = -\dfrac{dn_l}{d\xi} - \dfrac{dn_{2h}}{d\xi} < 0$;

(d) $\dfrac{dR_l}{d\xi} = \dfrac{-E_{lu} A_l A_3 q_l}{\det_\xi} > 0$; (e) $\dfrac{du}{d\xi} = \dfrac{A_l A_3 q_l^2}{\det_\xi} < 0$; (f) $\dfrac{dH_2}{d\xi} = \dfrac{d\left(n_2 q \left(R_a, u \right) \right)}{d\xi} = \dfrac{dn_2}{d\xi} q \left(R_a, u \right) + n_2 q_{2u} \dfrac{du}{d\xi} < 0$;

(g) $\dfrac{dD_h}{d\xi} = \dfrac{1}{w} \dfrac{dn_{2h}}{d\xi} > 0$; (h) $\dfrac{dD_r}{d\xi} = \delta \left(\sqrt{2\delta\psi T_r n_{2r} / \xi} + \sqrt{2\delta\psi \xi T_r / n_{2r}} \dfrac{dn_{2r}}{d\xi} \right) > 0$, due to

$$\dfrac{dTC_r}{d\xi} = 2\kappa \left(\sqrt{2\delta\psi T_r n_{2r} / \xi} + \sqrt{2\delta\psi \xi T_r / n_{2r}} \dfrac{dn_{2r}}{d\xi} \right) = \dfrac{dTC_h}{d\xi} = A_l \dfrac{dn_{2h}}{d\xi} > 0 . \;\square$$

Appendix B: Proof of Proposition 2
Proof: Similar to (A) but totally differentiate Eqs. (15), (16), (19) and (13) wrt ϕ .

The determinant is negative: $\det_\phi = E_{1u} A_1 A_2 n_l q_{R_l} - E_{2u} \left(A_1 + A_2 \right) q_l^2 - A_1 A_2 q_{lu} H_1 < 0$. Then, we find:

(a) $\dfrac{dn_l}{d\phi} = \dfrac{-A_l n_l \left(E_{1u} q_{R_l} - q_{lu} q_l \right)}{\det_\phi} < 0$; (b) $\dfrac{dn_{2h}}{d\phi} = \dfrac{E_{2u} q_l^2}{\det_\phi} < 0$; (c) $\dfrac{dn_{2r}}{d\phi} = -\dfrac{dn_l}{d\phi} - \dfrac{dn_{2h}}{d\phi} > 0$;

(d) $\dfrac{dR_l}{d\phi} = \dfrac{E_{1u} A_l q_l}{\det_\phi} < 0$; (e) $\dfrac{du}{d\phi} = \dfrac{-A_l q_l^2}{\det_\phi} > 0$; (f) $\dfrac{dH_2}{d\phi} = \dfrac{d\left(n_2 q \left(R_a, u \right) \right)}{d\phi} = \dfrac{dn_2}{d\phi} q_2 + n_2 q_{2u} \dfrac{du}{d\phi} > 0$;

(g) $\dfrac{dD_h}{d\phi} = \dfrac{1}{w} \dfrac{dn_{2h}}{d\phi} < 0$; (h) $\dfrac{dD_r}{d\phi} = \delta \sqrt{2\delta\psi \xi T_r / n_{2r}} \dfrac{dn_{2r}}{d\phi} > 0 . \;\square$

Appendix C: Proof of Lemma 1
Proof: Totally differentiating Eqs. (13), and (15) getting

$q \left(R_l, u \right) dn_l + n_l q_{R_l} dR_l + n_l q_{lu} du = 0$;

$q \left(R_l, u \right) dR_l + E_{1u} du = 0$.

Combining these two equations, getting

$$dn_l = \left(n_l q_{R_l} \dfrac{E_{1u}}{q \left(R_l, u \right)} - n_l q_{lu} \right) du \Big/ q \left(R_l, u \right) .$$

Under the Pareto-improving transit scheme, there has $u^p > u^e$, i.e., $du > 0$, then we obtain $dn_l < 0$. $\square$

Appendix D: Proof of Lemma 2 and 3

Proof: (i) First, we prove the uniqueness of $\underline{\xi}$ or $\bar{\sigma}$. For a given H_2 , from Eq. (14), $n_2 = H_2 / q \left(R_a + s, u \right)$. We plug this into Eq. (13), getting $Nq \left(R_l, u \right) q \left(R_a + s, u \right) - H_1 q \left(R_a + s, u \right) - H_2 q \left(R_l, u \right) = 0$. Totally differentiating it and Eq. (15), getting

$dR_l / du = - \left(n_l q \left(R_a + s, u \right) q_{lu} + n_2 q \left(R_l, u \right) q_{2u} \right) \big/ n_l q \left(R_a + s, u \right) q_{R_l} > 0$; $dR_l / du = -E_{1u} / q \left(R_l, u \right) < 0$.

Hence, the two equations can intersect once in R_l, u , and the solution R_l, u are uniquely found for given H_2 and s . Then, from Eq. (13), n_l and from Eq. (16), n_{2h} are uniquely found. According to population conservation, n_{2r} is uniquely found. Finally, from Eq. (19), $\underline{\xi}$ or $\bar{\sigma}$ is uniquely found.

(ii) Now, we prove the uniqueness of $\bar{\xi}$ or $\underline{\sigma}$. For a given u , due to the fact that $E \left(R_l, u \right)$ decreases

monotonically in R_l, the value of R_l is uniquely determined by Eq. (15). Next, from (13), $n_l = H_l/q(R_l, u)$ is unique and from Eq. (16), n_{2h} is unique because R_l, u have been found uniquely. According to population conservation, n_{2r} is uniquely found. Finally, from Eq. (19), $\bar{\xi}$ or $\underline{\sigma}$ is uniquely found. $\square$

Appendix E: Proof of Proposition 6

Proof: (i) Keeping u constant and setting $\phi = 0$, following (A1) but totally differentiate Eqs. (15), (16), (19) and (13) wrt s. The determinant is negative: $\det_{s\bar{\xi}} = -A_1 A_3 q_l^2 < 0$, and $\dfrac{d\bar{\xi}}{ds} = \dfrac{q_l^2 q_2 (A_1 + A_2)}{\det_{s\bar{\xi}}} < 0$. $\square$

(ii) With instrument $\{s, \underline{\xi}\}$, H_2 is unchanged, thus $dH_2/ds = 0$.

Following (A1), differentiate Eqs. (15), (16), (19) and (13) totally wrt s.

The determinant is negative: $\det_{s\underline{\xi}} = -\dfrac{A_1 A_3 \left(n_2 q_{2u} q_l^2 + n_1 q_1 q_2 q_{1u} - E_{1u} n_1 q_2 q_{R_l} \right)}{n_2 q_{2u}} < 0$. Then, we have

$$\frac{d\underline{\xi}}{ds} = \frac{\left(A_1 A_3 n_1 n_2 q_{R_2} \left(E_{1u} q_{R_l} - q_l q_{1u} \right) - (A_1 + A_2)\left(E_{2u} n_2 q_l^2 q_{R_2} + E_{1u} n_1 q_2^2 q_{R_l} - n q_1 q_2^2 q_{1u} \right) \right) / (n_2 q_{2u})}{\det_{s\underline{\xi}}} < 0. \square$$

Appendix F: Proof of Proposition 7

Proof: Keeping s constant and setting $\phi = \sigma n_2 q_2 s / (2\kappa n_{2r})$, differentiate Eqs. (15), (16), (19) and (13) totally with σ. The determinant is

$$\det_{\sigma} = -\frac{q_l^2 \left(E_{2u} \sigma n_2 q_2 s + E_{2u} A_2 n_{2r}^2 + E_{2u} A_1 n_{2r}^2 - q_{2u} A_1 \sigma n_2 s n_{2r} \right) + A_1 n_1 \left(q_{1u} q_l - E_{1u} q_{R_l} \right)\left(A_2 n_{2r}^2 + \sigma n_{2h} q_2 s \right)}{n_{2r}^2}.$$

The sign of $\det_{\sigma}$ is negative, evaluated at $\sigma = 0$. Then, using Cramer's rule, we find:

$$\frac{du}{d\sigma} = \frac{-A_1 q_l^2 q_2 s n_2 / n_{2r}}{\det_{\sigma}} > 0 \; ; \frac{dH_2}{d\sigma} = \frac{dn_2}{d\sigma} q_2 + n_2 q_{2u} \frac{du}{d\sigma} > 0 \text{ as } \frac{dn_1}{d\sigma} = \frac{-A_1 n_1 q_2 s n_2 \left(E_{1u} q_{R_l} - h_{1u} q_l \right) / n_{2r}}{\det_{\sigma}} < 0. \square$$

References

Alonso, W., 1964. Location and Land Use: Toward a General Theory of Land Rent. Harvard University Press, Cambridge, Massachusetts.

Amirgholy, M., Gonzales, E.J. 2016. Demand responsive transit systems with time-dependent demand: User equilibrium, system optimum, and management strategy. Transportation Research Part B 92, 234-252.

Anas, A., Pines, D., 2008. Anti-sprawl policies in a system of congested cities. Regional Science and Urban Economics 38, 408-423.

Anas, A., Xu, R., 1999. Congestion, land use, and job dispersion: A general equilibrium model. Journal of Urban Economics 45 (3), 451-473.

Arnott, R., 1998. Congestion tolling and urban spatial structure. Journal of Regional Science 38 (3), 495-504.

Arnott, R., de Palma, A., Lindsey, R., 1990. Departure time and route choice for the morning commute. Transportation Research Part B 24 (3), 209-228.

Boyce, D.E., Southworth, F., 1979. Quasi-dynamic urban-location models with endogenously determined travel costs. Environment and Planning A 11, 575–584.

Brabec, E., Smith, C., 2002. Agricultural land fragmentation: the spatial effects of three land protection strategies in the eastern United States. Landscape and Urban Planning 58, 255-268.

Brueckner, J.K., 2007. Urban growth boundaries: An effective second-best remedy for unpriced traffic congestion? Journal of Housing Economics 16, 263–273.

Buyukeren, A.C., Hiramatsu, T., 2016. Anti-congestion policies in cities with public transportation. Journal of Economic Geography 16 (2), 395-421.

Capozza, D.R., 1973. Subways and land use. Environment Planning 5, 555–576.

Chen, Y.J., Li, Z.C., Lam, W.H.K., Choi, K., 2016. Tradable location tax credit scheme for balancing traffic congestion and environmental externalities. International Journal of Sustainable Transportation 10 (10), 917-934.

Cohen, S.E., 1994. Greenbelts in London and Jerusalem. Geographical Review 84 (1), 74-89.

Gonzales, E.J., Daganzo, C.F., 2012. Morning commute with competing modes and distributed demand: User equilibrium, system optimum, and pricing. Transportation Research Part B 46 (10), 1519-1534.

Gubins, S., Verhoef, E.T., 2014. Dynamic bottleneck congestion and residential land use in the monocentric city. Journal of Urban Economics 80, 51-61.

Haring, J.E., Slobko, T., Chapman, J., 1976. The impact of alternative transportation systems on urban structure. Journal of Urban Economics 3 (1), 14-30.

Jun, M.J., 2004. The effects of Portland's urban growth boundary on urban development patterns and commuting. Urban Studies 41 (7), 1333-1348.

Li, Z.C., Lam, W.H.K., Wong, S.C., Sumalee, A., 2012. Design of a rail transit line for profit maximization in a linear transportation corridor. Transportation Research Part E 48 (1), 50-70.

Li, Z.C., Peng, Y.T., 2016. Modeling the effects of vehicle emission taxes on residential location choices of different-income households. Transportation Research Part D 48, 248-266.

Ma, X.S., Lo, H.K., 2012. Modeling transport management and land use over time. Transportation Research Part B 46 (6), 687-709.

McDonald, J.F., Osuji, C.I., 1995. The effect of anticipated transportation improvement on residential land values. Regional Science and Urban Economics 25 (3), 261-278.

Mills, E.S., 1967. An aggregative model of resource allocation in a metropolitan area. American Economic Review 57 (2), 197–210.

Muth, R.F., 1969. Cities and Housing. University of Chicago Press, Chicago.

Pines, D., Sadka, E., 1985. Zoning, first-best, second-best, and third-best criteria for allocating land for roads. Journal of Urban Economics 17 (2): 167-183.

Sasaki, K., 1989. Transportation system change and urban structure in two-transport mode setting. Journal of Urban Economics 25 (3), 346–367.

Small, K. A., 1982. The scheduling of consumer activities: work trips. American Economic Review 72 (3), 467–479.

Solow, R.M., 1972. Congestion, density and the use of land in transportation. The Swedish Journal of Economics 74 (1), 161-173.

Strotz, R.H., 1965. Urban transportation parables, in 'The Public Economy of Urban Communities' (J. Margolis Ed.), Johns Hopkins University Press, Baltimore, 127–169.

Su, Q., DeSalvo, J.S., 2008. The effect of transportation subsidies on urban sprawl. Journal of Regional Science 48 (3), 567–594.

Takayama, Y., Kuwahara, M., 2016. Bottleneck congestion and residential location of heterogeneous commuters. https://mpra.ub.uni-muenchen.de/68940/.

Verhoef, E.T., 2005. Second-best congestion pricing schemes in the monocentric city. Journal of Urban Economics 58 (3), 367-388.

Verhoef, E.T., Nijkamp, P., 2008. Urban environmental externalities, agglomeration forces, and the technological 'deus ex machina'. Environment and Planning A 40 (4), 928-947.

Vickrey, W.S., 1969. Congestion theory and transport investment. American Economic Review 59 (2), 251–261.

Wheaton, W.C., 1998. Land use and density in cities with congestion. Journal of Urban Economics 43 (2): 258-272.

Wu, W.X., Huang, H.J., 2014. Equilibrium and modal split in a competitive highway/transit under different road-use pricing strategies. Journal of Transport Economics and Policy 48 (1), 153-169

Xiao, L.L., Huang, H.J., Liu, R., 2015a. Congestion behaviour and tolls in a bottleneck model with stochastic capacity. Transportation Science 49 (1), 46-55.

Xiao, L.L, Huang, H.J., Liu, R., 2015b. Tradable credit scheme for rush house travel choice with heterogeneous commuters. Advances in Mechanical Engineering 7 (1), 1-12.

Zhao, F., Sun, H., Wu, J., Gao, Z. and Liu, R., 2016. Analysis of road network pattern considering population distribution and central business district. PLoS ONE 11(3), e0151676.